

# Sprint – 1

TeamID - PNT2022TMID10506

Date : 29-Oct-2022

## Sprint - 1

### Machine Learning based Vehicle Performance Analyzer

#### Description

**Goal:** This sprint delivers a feature so that we can install the pre-requisites, collecting the data, importing libraries, read dataset and clean dataset.

**Method:** Collecting data for training the ML model is the basic step in the learning pipeline. The predictions made by ML systems can only be as good as the data on which they have been trained. Real-world raw data and images are often incomplete, inconsistent, and lacking in certain behaviors or trends. They are also likely to contain many errors. So, once collected, they are pre-processed into a format the machine learning algorithm can use for the model.

#### **Metrics :**

- Pre-cleaned, freely available datasets. If the problem statement (for example, image classification, object recognition) aligns with a clean, pre-existing, properly formulated dataset, then take advantage of existing, open-source expertise.
- Data integration. Combining multiple datasets to get a large corpus can overcome incompleteness in a single dataset.
- Data normalization. The size of a dataset affects the memory and processing required for iterations during training. Normalization reduces the size by reducing the order and magnitude of data.

## SPRINT - 1

